Received December 12, 2019, accepted January 1, 2020, date of publication January 9, 2020, date of current version January 22, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2965462

Leveraging Artificial Intelligence for Large-Scale Plant Phenology Studies From Noisy Time-Lapse Images

DAVID L. P. CORREIA[1], WASSIM BOUACHIR[2], DAVID GERVAIS[1], DEEPA PURESWARAN[1], DANIEL D. KNEESHAW[3], AND LOUIS DE GRANDPRÉ[1]

[1]Laurentian Forestry Centre, Canadian Forest Service, Natural Resources Canada, Québec, QC G1V 4C7, Canada
[2]Department of Science and Technology, TÉLUQ University, Montréal, QC H2S 3L5, Canada
[3]Centre d'étude de la forêt (CEF), Université du Québec à Montréal, Montréal, QC H3C 3P8, Canada

Corresponding author: Louis De Grandpré (louis.degrandpre@canada.ca)

This work was supported in part by the Natural Resources Canada ADM Innovation Fund, and in part by the Canadian Foundation for Innovation (CFI) through the Project SmartForests Canada: A network of monitoring plots and plantations for modeling and adapting forests to climate change.

ABSTRACT Phenology has become a field of growing importance due to the increasingly apparent impacts of climate change. However, the time-consuming, subjective and tedious nature of traditional human field observations have hindered the development of large-scale phenology networks. Such networks are rare and rely on time-lapse cameras and simplistic color indexes to monitor phenology. To automatize rapid, detailed and repeatable analyzes, we propose an Artificial Intelligence (AI) framework based on machine learning and computer vision techniques. Our approach extracts multiple ecologically-relevant indicators from time-lapse digital photography datasets. The proposed framework consists of three main components: (i) a random forest model to automatically select relevant images based on color information; (ii) a convolutional neural network (CNN) to identify and localize open tree buds; and (iii) a density-based spatial clustering algorithm to cluster open bud detections across the time-series. We tested this framework on a dataset including thousands of black spruce and balsam fir tree images captured using our phenological camera network. The performed experiments showed the efficiency of the proposed approach under challenging perturbation factors, such as significant image noise. Our framework is exceedingly faster and more accurate than human analysts, reducing the time-series processing time from multiple days to under an hour. The proposed methodology is particularly appropriate for large-scale and long-term analyzes of ecological imagery datasets. Our work demonstrates that the use of computer vision and machine learning methods represents a promising direction for the implementation of national, continental, or even global plant phenology networks.

INDEX TERMS Balsam fir, black spruce, computer vision, convolutional neural network, deep learning, forest ecology, object detection, tree budburst.

I. INTRODUCTION

Recent impacts of climate change on plant and animal life cycle events (i.e. phenology), along with their potential cascading effects on ecosystem functioning, have cemented this field of study as a crucial component of global change science [1]. Plant phenological shifts have recently been found to influence the water cycle through impacts on evapotranspiration [2], whereas the observed alteration of the timing of phenological events is likely to lead to the desynchronization of plant-animal interactions [3]. This decoupling of trophic interactions through changes in species phenology has the potential to induce long-lasting effects on biodiversity.

Phenological studies have been traditionally undertaken with ground-based measurements. These measurements are very useful, but conventional sampling methodologies are exceedingly slow. For instance, in [4], the timing of budburst was monitored by returning to the same sites every 2-3 days and manually examining the buds on up to 50 branches per tree. Such traditional approaches render large-scale

The associate editor coordinating the review of this manuscript and approving it for publication was Shadi Alawneh.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see http://creativecommons.org/licenses/by/4.0/

studies prohibitively time-consuming and expensive. An alternative is the development of large-scale phenology networks using inexpensive time-lapse cameras (e.g. PhenoCam in the US [5], Phenological Eyes Network in Japan [6]).

The implementation of such large-scale phenology networks results in an interesting paradigm shift: viability no longer depends on data acquisition, which is relatively cheap. Instead, project feasibility is greatly dependent on the development of tools capable of efficiently analyzing thousands of images and terabytes of data. These analytical tools are generally based on color analysis of red, green and blue bands (e.g. [5], [7]). They are thus capable of measuring indices such as landscape greening but are currently unable to extract more sophisticated phenological data, such as the number of buds open in a given tree, which are directly linked to individual tree growth and herbivore food availability. Recent advances in computer vision and the democratization of neural network models have given researchers the tools necessary to exploit the analytical potential of these large-scale phenology datasets.

Convolutional neural networks (CNNs) have revolutionized the field of computer vision. CNNs now perform nearly as well as humans in object classification tasks (i.e. identifying and classifying objects in an image [8]) and are performing remarkably well in spatially assigning objects (i.e. where are objects located in an image [9]). However, they remain vastly underutilized in ecological studies. In animal ecology, a few researchers have recently started using deep learning methods to process camera-trap data [10], while [11] proposed a methodology to automatically detect turtles in drone imagery using CNNs. In plant ecology, existing neural network applications are scarce and include the identification of multiple plant species from images of collected leaves in white backgrounds [12], the mapping of forest cover type and structure from aerial and satellite imagery [13], and the development of a CNN-based decision support system to help forest managers estimate the number of planting microsites on planting blocks [14].

In phenology, studies leveraging the power of CNNs and digital repeat photography are even rarer. In a literature review on this subject, only two studies were found: [15], which fine-tuned a CNN to classify phenological states of agricultural plants, and [16], which employed a CNN to identify the occurrence of snow in thousands of near-surface images from the PhenoCam network. Both of these studies relied only on fine-tuning CNNs capable of identifying the presence of a particular object in an image. Considering that the great potential of object detectors remains unexplored in phenology studies, there is a compelling need to develop methodologies capable of automating the analysis of large datasets for large-scale longitudinal studies.

In this work, we developed a 3-step artificial intelligence framework capable of exploiting the synergies between cutting-edge, CNN-based object detectors, traditional machine learning algorithms and longitudinally repeated digital photography to extract more detailed data from past and future ground and near-surface phenology imagery datasets. First, our framework automatizes image selection with a machine learning algorithm in order to reduce the amount of noise inherently present in time-lapse datasets. Second, it identifies the presence and location of multiple open buds in each image with a CNN. Finally, our data processing pipeline uses a clustering algorithm to summarize individual open bud detections into meaningful phenological measurements, namely an estimate of the total number of open buds, the proportion of open buds per day, the date of budburst onset, and the rate of budburst. To our knowledge, our work is the first to propose an AI approach capable of identifying and localizing structures of phenological interest. The obtained measurements are significantly more comprehensive than the color indices traditionally used. The proposed framework was implemented in a coniferous forest site in the North Shore of Quebec, Canada,[1] and is shown to be applicable to numerous plant communities worldwide.

II. MATERIALS AND METHODS

A. STUDY AREA

Time-lapse digital cameras were installed in 2014 in three sites of the North Shore region of Quebec, Canada (Fig. 1). These sites follow a latitudinal gradient of black spruce (*Picea mariana*) and balsam fir (*Abies balsamifera*) mixed stands. In this region, tree budburst typically occurs between the end of May and the beginning of June, with balsam fir usually undergoing budburst a week or two before black spruce [17].

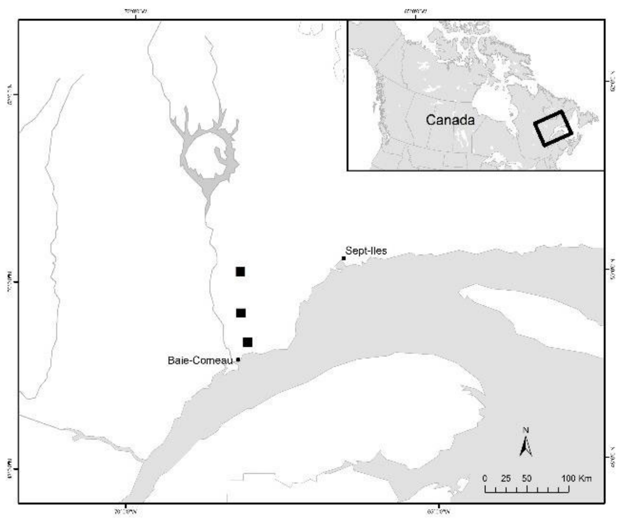

FIGURE 1. Map of the study area.

B. TIME-LAPSE IMAGERY AND COLLECTION

SpyPoint TinyHD 8 megapixel cameras (www.spypoint.com) were installed every year before the end of April and were removed in August (Fig. 2). Hence, all cameras were installed well before tree budburst and were removed weeks after

[1] BudCam project: https://apps-scf-cfs.rncan.gc.ca/budcam/en

FIGURE 2. Example of an installed SpyPoint TinyHD camera.

budburst onset took place. All cameras were set up to take RGB (Red, Green and Blue) images every 30 minutes from 5.00 AM to 20.00 PM, local time. To ensure that most buds were recorded, trees that were 2 to 4 meters tall were targeted. Since all cameras were strapped onto nearby trees, the distance to the target tree was not fixed, but was usually around 5 meters.

C. AUTOMATIC IMAGE SELECTION

A considerable proportion of the images available in each time-series were inadequate for bud detection because they were either too bright due to direct sunlight, too dark due to a lack of sunlight, or too misty.

Hence, a random forest model [19] was trained to automatically remove such unusable images from the analysis based on color features (Fig. 3). The random forest model was trained on 4102 high-quality images where trees were clearly visible and on 3701 low-quality images where they were not. This dataset was manually collated from the available time-lapse images. Each image was decomposed into RGB (Red, Green, Blue) and HSV (Hue, Saturation, Value) color histograms with 32 bins per channel (except hue, which had 16 bins). The random forest model was trained with the *scikit-learn* python module (version 0.20.3; [20]). The model had 250 trees, a minimum number of one sample per leaf, used 13 features per split (i.e., the square root of all available features) and was grown via entropy. To improve prediction of budburst phenology, we only used up to 10 images per day. If more than 10 high-quality images were available, we used the ones that had the best random forest prediction probability scores.

D. CONVOLUTIONAL NEURAL NETWORK

1) IMAGE ANNOTATION

For training CNN models, we started by manually determining budburst onset for all train and test trees (33 black spruce and 29 balsam fir) Open buds were manually annotated (VIA annotation tool; [18]) for 15 black spruce and 13 balsam fir trees, whereas closed buds were annotated for 13 black spruce and 10 balsam fir trees. Usually, there were 30 images available per tree per bud type (open and closed). The open bud images were chosen at the beginning of the manually determined budburst onset period. Evergreen buds stay dormant throughout winter, at which time they are small, brown, rounded structures. Buds usually start opening in late spring, which results in a gradual increase in size and shift in color towards a bright green, until the new leaves are completely unfolded. Buds were only classified as "open" at the beginning of the phenophase shift, when their color started shifting from light brown to light green (Fig. 4). As buds were approximately 10 to 30 pixels long, bounding boxes were 38 × 38-pixel squares centered on the corresponding bud. Closed buds were annotated and passed onto the CNN as hard backgrounds because they are visually quite similar to the initial stages of open buds and could be easily misclassified by the CNN. Since balsam fir and black spruce are evergreen, buds were hard to locate among the previous year's foliage. The relatively low image quality and the large potential number of buds per image made it impossible to annotate all buds in each image. Additionally, a marginal number of annotated open buds is unlikely to correspond to true open buds.

Due to the small size of the objects of interest and the large size of the images (3264 × 2448 pixels), each image was split into multiple tiles of 200 pixels, which overlapped by 40 pixels to ensure that buds located close to tile edges were also identified by the CNN (Fig. 3). In total, 21 172 black spruce open bud annotations in 10 374 tiles of 447 tree images were used to train and validate the CNN, along with 11 188 hard background tiles (including 6743 tiles with closed buds) of 1217 images of hard backgrounds. For the balsam fir CNN, 26 435 open bud annotations in 14 341 tiles of 638 tree images were used during training and validation, in addition to 15 971 tiles (including 9832 closed bud tiles) of 1093 images of hard backgrounds. The validation datasets included open bud annotation data from 3 out of 15 black spruces and 2 out of 13 balsam firs, as well as additional hard background data from 3 black spruces and 2 balsam firs. The validation datasets consisted of 22.5% of the black spruce tiles and 17.5% of the balsam fir tiles available. Each time an image was passed to the CNN, it was rescaled to 500 × 500 pixels and submitted to multiple random data augmentation techniques, namely horizontal and vertical flipping, rotation, translation, shearing and rescaling (Fig. 3).

2) ARCHITECTURE AND TRAINING

The CNN we used was RetinaNet, a state-of-the-art object detector developed by the Facebook AI Research group [9]. This one-stage detector was chosen for three main reasons: it is precise, fast, and it is readily available (https://github.com/fizyr/keras-retinanet). RetinaNet was trained with the Keras neural network library (version 2.2.4; [21]) and Tensorflow (version 1.12; [22]) on a NVIDIA GeForce RTX 2070 graphics processing unit. RetinaNet incorporates two major improvements over other one-stage detectors: (i) a focal loss function; and (ii) feature pyramid networks (FPNs). The focal loss function is an adaptation

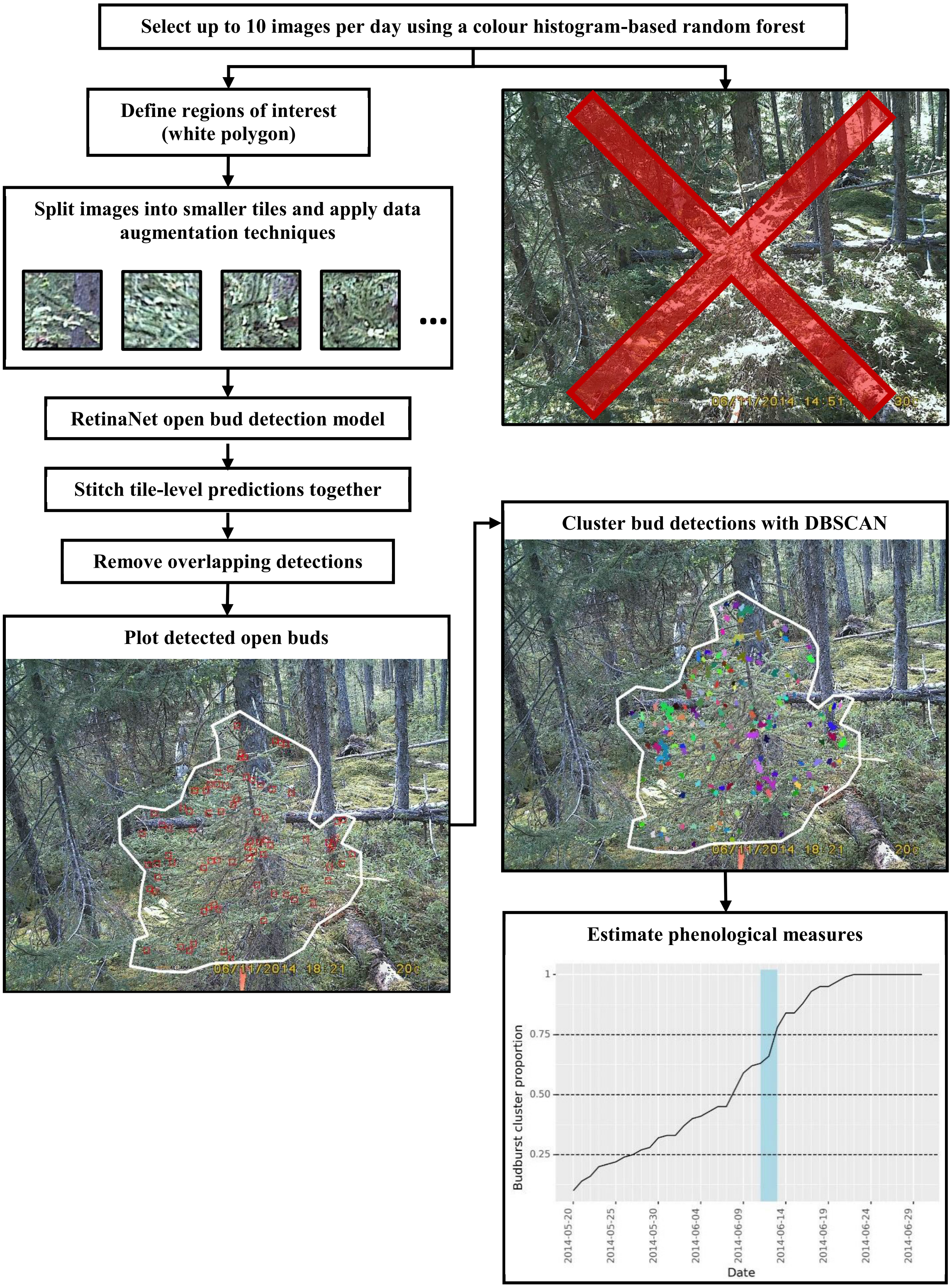

FIGURE 3. Detailed overview of the application of the 3-step framework that estimates phenological measures from time-lapse images. White polygons in the bottom left and middle right figures define the regions of interest analyzed. Red squares in the bottom left figure are bud detections. Open bud clusters in the middle right figure are randomly colored. The blue band in the bottom right figure corresponds to the manually determined period of budburst.

of the commonly used cross entropy loss that addresses the extreme class imbalance between background and foreground classes faced by predictors. This loss function greatly down-weights the importance of easily classified examples and focuses learning on hard misclassified examples [9]. The second improvement, FPNs, takes advantage of the

FIGURE 4. Example of 200 × 200 pixel tiles with closed (left) and open (right) balsam fir (top) and black spruce (bottom) buds. Buds are identified by red circles.

hierarchical structure of CNNs to generate multiple feature maps at different scales. FPNs are thus capable of augmenting the number of feature maps available and improving multi-scale predictions [23].

RetinaNet is composed of one backbone network and two task-specific subnetworks. We used a ResNet152 backbone [24], whereas we kept the same subnetworks described in the original paper [9]. Other backbones were tested, namely ResNet50, ResNet101, VGG16, VGG19 and Densenet121, but ResNet152 was the one that performed best (comparison results are not shown). The two subnetworks previously mentioned use the convolutional feature maps generated by the backbone neural network and the FPN to perform object classification and bounding box regression.

We started training our black spruce RetinaNet model with pre-trained COCO weights (Common Objects in Context; http://cocodataset.org), which provided a much better starting point than randomized weight initialization. We trained our model for up to 20 epochs with an ADAM optimizer [25] with an initial learning rate of $1e^{-5}$. The balsam fir RetinaNet was initialized with the black spruce model weights. These weights provided a better starting point than the pre-trained COCO weights, probably because the black spruce RetinaNet had already learned to identify similar buds amongst tree foliage. The balsam fir CNN was trained with an initial learning rate of $1e^{-5}$ and an ADAM optimizer for a maximum of 60 epochs. Only the CNNs that had the highest average precision rate were kept.

E. BUD DETECTION POST-PROCESSING AND CLUSTERING

For each time-series, a polygonal region of interest (ROI) was manually defined around the tree of interest (Fig. 3). Buds were only considered if they were inside the ROI, which reduced the proportion of false positives and ensured that all detected buds belonged to the same tree. Since the cameras were fixed throughout each season, a single ROI mask could be used for the whole time-series, as long as the ROI was large enough to account for daily tree branch movements (e.g. due to wind).

Overlapping detections were trimmed via non-maximum suppression using an intersection over union (IOU) value of 0.2. Hence, if the area of the bounding box defined by two detections overlapped by more than 20%, we discarded the detection with the lowest RetinaNet prediction probability.

Depending on image quality, the number of buds detected per image could vary considerably between images taken within the same day. Additionally, the CNN was incapable of estimating the proportion of open buds at a given date because it was unable to re-identify the same bud in different images. In order to overcome this limitation, we used DBSCAN (density-based spatial clustering of applications with noise; [26]) to cluster bud detections. This algorithm groups together points that are closely packed and classifies points in low-density regions as outliers. DBSCAN is suitable for this clustering problem because it allows the user to define a maximum distance between points in the same cluster and a minimum number of points per group. DBSCAN was initially used to cluster day-level bud predictions using a maximum distance radius of 25 pixels and a minimum number of 3 detections per group. Then, DBSCAN was used again to regroup all daily bud detection clusters using a maximum distance radius of 10 pixels and a minimum number of 4 elements per group. Clustering parameters were fine-tuned to reduce the number of false positives on the day-level clustering phase and prevent the formation of large clusters on the site-level clustering phase.

III. RESULTS

All analyzes were performed with Python (version 3.6.8; www.python.org). Data processing and handling was done with the Pandas (version 0.24.2; [27]) and Numpy (version 1.16.2; [28]) libraries. Images were processed with the OpenCV library (version 4.0.; [29]) and figures were plotted with the Plotnine library (version 0.5.1; http://plotnine.readthedocs.io).

A. AUTOMATIC IMAGE SELECTION

The random forest developed to automatize selection of high-quality images was evaluated according to validation precision and recall. Precision, which is calculated as the number of true positives divided by the sum of true positives and false positives, describes the ability of the model to identify only images of interest (i.e. annotated positives). Recall, defined as the number of true positives divided by the sum of true positives and false negatives, represents the capacity of the model to detect all points of interest. These two metrics can be formulated as follows:

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives};$$
$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives};$$

Our random forest model was very effective, as it achieved a precision of 96.74% for high-quality images and 89.9% for low-quality images, and a recall rate of 89.64% for high-quality images and 96.82% for low-quality images (Table 1). Manual revision of the images selected by the model confirmed the efficiency of this random forest model.

TABLE 1. Confusion matrix of color histogram-based random forest model.

Observed	Predicted	
	Low-quality	High-quality
Low-quality	1219 (96.8%)	40 (3.2%)
High-quality	137 (10.4%)	1186 (89.6%)

B. HYPERPARAMETER FINE-TUNING

We tested several hyperparameters during convolutional neural network (CNN) training other than multiple backbones. We cumulatively froze all ResNet152 convolutional layer blocks (e.g. the first block, the first and second block, etc...), but concluded that not freezing layers yielded the best results. This is likely because the objects we were trying to detect were very different from the objects present in the original COCO dataset used to initialize model training. We tested multiple bounding box sizes, from 30 × 30 pixels to 50 × 50 pixels and found that 38 × 38 pixels boxes resulted in models with higher precision and recall. This is likely because 38 × 38-pixel boxes are large enough to include a small margin around the target bud for context, without letting the noise present in the margin overwhelm the neural network. We also tested several tile resizing sizes (300 to 800 pixels) and found that a 500-pixel resizing parameter resulted in more precise models, presumably because of a trade-off between greater bud size and lower image resolution. Finally, we evaluated several class imbalance ratios and kept a 1 to 1 tile ratio between open bud and hard background tiles, which resulted in the best compromise between too much noise and not enough data.

C. RETINANET MODEL PERFORMANCE

RetinaNet model performance was evaluated using validation precision and recall. After training, the models were tuned to favor recall over precision because a considerable proportion of open buds were not annotated, meaning that a large number of false positives actually corresponded to true, non-annotated positives. The annotation process in itself is extremely time consuming and visually strenuous for observers because the annotated objects are numerous, but the objects themselves are very small (10 to 30 pixels long). With a detection probability threshold of 0.2, the black spruce model achieved 82.53% recall and 14.97% precision on the corresponding validation dataset, whereas the balsam fir model achieved 52.12% validation recall and 21.83% validation precision (Fig. 5).

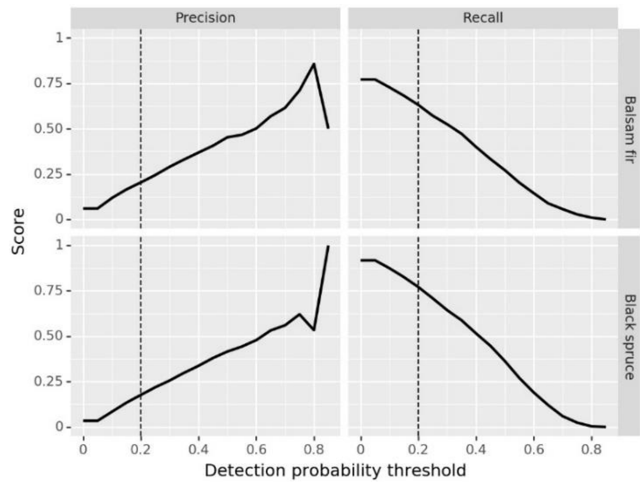

FIGURE 5. Balsam fir (top) and black spruce (bottom) RetinaNet model validation precision (left) and recall (right) scores. These scores are presented according to increasing detection probability thresholds. The vertical dashed line represents the chosen probability detection threshold.

D. BUDBURST ONSET PREDICTION

Besides using the traditional validation metrics mentioned in the previous section, we validated our models by comparing neural network predicted budburst onset dates with manually estimated dates from test sites not used in model training (Fig. 6). The number of open buds detected was calculated for a maximum of 10 images per day, depending on the number of images selected by the color histogram-based random forest model. The neural network open bud cluster detection increase coincided with the budburst onset dates that had been manually identified for both black spruce and balsam fir (Fig. 6).

IV. DISCUSSION

In this study, we present the first example of an AI framework capable of identifying phenological activity in plants. Manually extracting this type of data from large digital repeat photography datasets is prohibitively time-consuming. However, our AI framework represents an exceptionally efficient alternative: it takes only a few seconds to identify and localize open buds in each image, while manual analysis takes several minutes per image. Furthermore, neural networks are capable of indefatigably analyzing data nonstop, 24 hours per day, whereas human analysts need to take frequent stops and are more prone to counting errors as weariness accumulates. The time it takes to extract phenological measures from an individual time-series is thus reduced from multiple days to less than an hour. Replicability of results is also better with convolutional neural networks than with human observers.

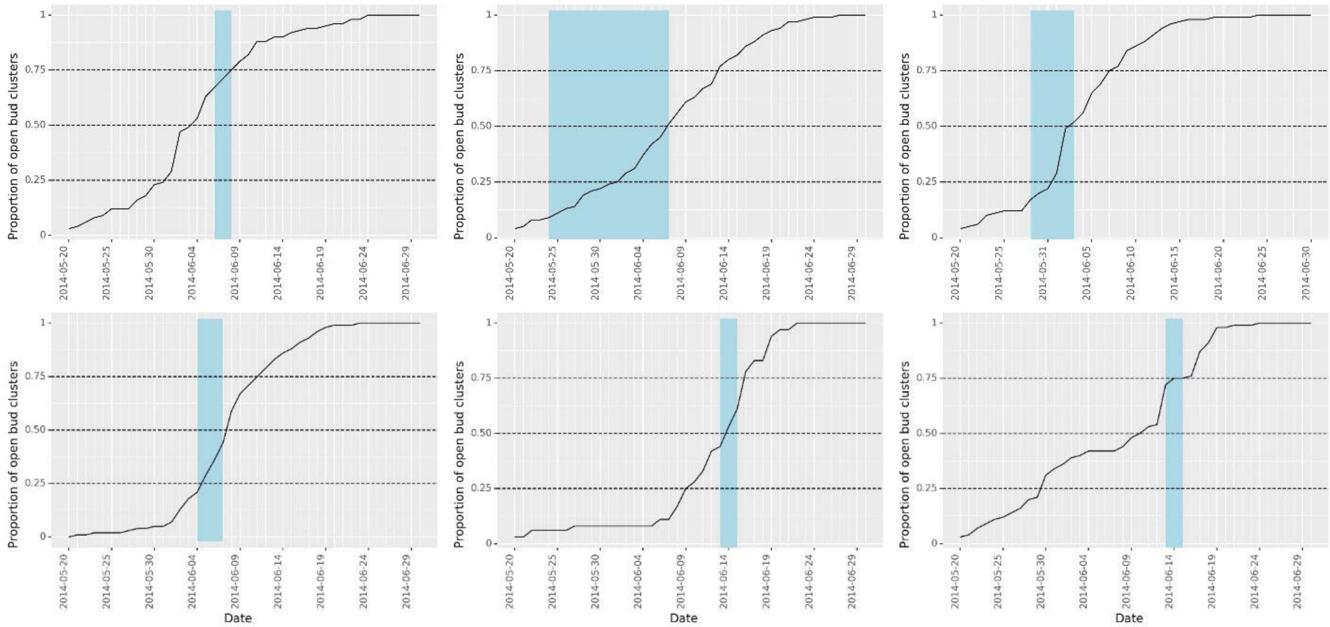

FIGURE 6. **Predicted cumulative proportion of open bud clusters per day of balsam fir (top) and black spruce (bottom) testing sites. The blue band represents the budburst onset period that was identified manually.**

Manually identifying open buds in time-lapse phenology images can actually be an onerous and challenging task. The trees examined in this study can easily have over a hundred buds, whereas older trees can have thousands to tens of thousands of buds. These structures are very small and are often similar to other features present in the same images, such as open sky patches and lichen spots. Additionally, the differences between closed buds and the initial stages of open buds can be somewhat subjective and can vary according to light conditions when the image was taken (Fig. 4). Due to these issues, our annotated training and validation datasets were quite noisy: many open buds were not annotated as such (Appendix S1), whereas some annotated open buds actually corresponded to closed buds or other alternative features. Nevertheless, our results reveal that RetinaNet model training is quite robust to noisy training data, in part because researchers can choose to prioritize either recall or precision.

Following initial analyzes, we chose to favor validation recall over validation precision. The former represents the ability of the model to detect all annotated objects in the dataset, while the latter quantifies the capacity of the model to identify only annotated objects. The choice of validation recall was made because a large number of open buds in the training and validation datasets were not annotated: this choice allowed us to identify most annotated buds, as well as multiple non-annotated buds within the validation dataset (Appendix S1). More importantly, our model correctly predicted the onset of budburst in the testing dataset (Fig. 6), which had not been used in model training and validation.

It is also relevant to point out that, as long as our object detector was able to identify enough buds to estimate budburst onset periods and rates, it did not need to achieve the extremely demanding precision and recall rates required in other fields (e.g. automated driving). Even though ecological datasets are often very noisy, we show that these novel deep learning approaches are capable of handling such data and we encourage researchers to experiment with them.

Ground-based phenological observations can have a considerable level of user-induced bias [30]. For instance, the phenophase shift assessed in our study can be subject to inter and intra-observer bias because it corresponds to a relatively slow, continuous change in bud color and size. Machine learning and deep learning algorithms offer us the opportunity to greatly reduce these biases: using a single model that always follows the same rules reduces inter and intra-observer bias. Such a reduction in observer bias represents an extraordinary opportunity for phenology citizen science-based projects, like the US National Phenology Network [31], to produce more complex and higher quality data. In these types of projects, citizen scientists are usually asked to describe phenological features, such as emerging leaves and flowering. While very useful, these measures are known to have a considerable level of observer bias due to their subjectivity, particularly during phenophase shifts [32]. If citizen scientists are tasked with uploading images instead of describing phenological features, the obvious synergies between the large amounts of data generated by citizen science projects and the data-hungry analytical power of artificial intelligence could be effectively exploited.

Crowdsourcing projects like CrowdCurio, which recruits citizen scientists to tag phenological structures in digitized herbarium specimen images [33], could also explore these synergies quite easily: citizen scientists annotate the data that allows researchers to train more efficient machine learning models, such as CNNs. CNNs also have the potential to revolutionize the type of analyzes done with large-scale phenology imagery datasets, such as PhenoCam [34]. Usually, these types of analyzes rely on RGB color indices, namely the green chromatic coordinate (e.g. [5], [35]). This index is easy to estimate and is correlated with important phenological events, such as canopy greening [36] and seasonal canopy-level photosynthesis [37]. However, the green chromatic coordinate is simultaneously affected by leaf color and canopy structure [35], is insensitive to significant levels of defoliation [38], and in coniferous-dominated forests is mostly associated with changes in pigmentation of existing leaves instead of leaf emergence and senescence [37]. By detecting the structures that these indexes represent, convolutional neural networks are capable of extracting new, and more complex data from existing datasets (e.g. rate of budburst instead of budburst onset; [39]) and greatly improve our understanding of plant phenology.

The 3-step framework presented in this study has delivered promising results, but represents only one of multiple possible ways artificial intelligence can be used to analyze digital repeat photography datasets. First, the same model can probably be applied to closely related species with similar bud structures (e.g. black and white spruce; *Picea glauca*). Alternatively, CNNs for new species can be fine-tuned with little data via transfer learning, where a model trained for a specific task is used as a starting point for a different model trained for a new task [40]. Second, CNNs could be trained to target distinct budding stages. Hence, instead of seeing a peak in the number of detected buds followed by a sharp decrease as buds develop, we would be able to follow new buds throughout their developmental process. Finally, other deep learning methods could be explored. The implementation of alternative RetinaNet backbones, such as ResNext [41] and Inception-Resnet v2 [42] could be assessed, or the performance of different models, such as light-head R-CNN [43], could be examined.

Computer vision is a rapidly expanding field of research and multiple computer scientists actively develop open-source implementations of cutting-edge algorithms (e.g. the RetinaNet algorithm implemented in this study: https://github.com/fizyr/keras-retinanet; Mask-RCNN: https://github.com/matterport/Mask_RCNN). These implementations are capable of efficiently extracting vast quantities of data and offer new and exciting opportunities across a wide range of fields, from camera trapping [10] to phenological analyzes, airborne tree species classification [44] and land cover classification [45]. Ecological sciences need to harness the potential of deep learning in order to develop novel ways of analyzing existing datasets, design innovative projects capable of producing more complex data and ultimately improve our understanding of ecosystem functions and processes.

V. CONCLUSION

The methodology proposed in this study shows what can be accomplished when artificial intelligence is used to process ecological data. Our framework is capable of quickly converting a time-lapse digital photography dataset into multiple ecologically-relevant indicators in three steps: (i) automatic selection of the most adequate images available in the time-series; (ii) identification and localization of multiple open buds per image; and (iii) clustering of bud detections for the estimation of various indicators of ecological importance. Our future work will focus both on the technological improvement of this framework and on the extension of our phenology networks to produce the larger amounts of data that we are now able to process.

REFERENCES

[1] S. Piao, Q. Liu, A. Chen, I. A. Janssens, Y. Fu, J. Dai, L. Liu, X. Lian, M. Shen, and X. Zhu, "Plant phenology and global climate change: Current progresses and challenges," *Glob. Change Biol.*, vol. 25, no. 6, pp. 1922–1940, Jun. 2019.

[2] J. H. Kim, T. Hwang, Y. Yang, C. L. Schaaf, E. Boose, and J. W. Munger, "Warming-induced earlier greenup leads to reduced stream discharge in a temperate mixed forest catchment," *J. Geophys. Res. Biogeosci.*, vol. 123, no. 6, pp. 1960–1975, Jun. 2018.

[3] H. M. Kharouba, J. Ehrlén, A. Gelman, K. Bolmgren, J. M. Allen, S. E. Travers, and E. M. Wolkovich, "Global shifts in the phenological synchrony of species interactions over recent decades," *Proc. Nat. Acad. Sci. USA*, vol. 115, no. 20, pp. 5211–5216, May 2018.

[4] D. S. Pureswaran, M. Neau, M. Marchand, L. De Grandpré, and D. Kneeshaw, "Phenological synchrony between eastern spruce budworm and its host trees increases with warmer temperatures in the boreal forest," *Ecol. Evol.*, vol. 9, no. 1, pp. 576–586, Jan. 2019.

[5] A. D. Richardson, K. Hufkens, T. Milliman, D. M. Aubrecht, M. Chen, J. M. Gray, M. R. Johnston, T. F. Keenan, S. T. Klosterman, M. Kosmala, E. K. Melaas, M. A. Friedl, and S. Frolking, "Tracking vegetation phenology across diverse North American biomes using PhenoCam imagery," *Sci. Data*, vol. 5, Mar. 2018, Art. no. 180028.

[6] K. N. Nasahara and S. Nagai, "Review: Development of an in situ observation network for terrestrial ecological remote sensing: The phenological eyes network (PEN)," *Ecol. Res.*, vol. 30, no. 2, pp. 211–223, Mar. 2015.

[7] J. L. O'Connell and M. Alber, "A smart classifier for extracting environmental data from digital image time-series: Applications for PhenoCam data in a tidal salt marsh," *Environ. Model. Softw.*, vol. 84, pp. 134–139, Oct. 2016.

[8] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015.

[9] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar, "Focal loss for dense object detection," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2017, pp. 2980–2988.

[10] S. Schneider, G. W. Taylor, S. Linquist, and S. C. Kremer, "Past, present and future approaches using computer vision for animal re-identification from camera trap data," *Methods Ecol. Evol.*, vol. 10, no. 4, pp. 461–470, Apr. 2019.

[11] Patrick C. Gray A. B. Fleishman, D. J. Klein, M. W. McKown, V. S. Bézy, K. J. Lohmann, and D. W. Johnston, "A convolutional neural network for detecting sea turtles in drone imagery," *Methods Ecol. Evol.*, vol. 10, pp. 345–355, Mar. 2019.

[12] P. Barré, B. C. Stöver, K. F. Müller, and V. Steinhage, "LeafNet: A computer vision system for automatic plant species identification," *Ecological Inform.*, vol. 40, pp. 50–56, Jul. 2017.

[13] T. Chang, B. Rasmussen, B. Dickson, and L. Zachmann, "Chimera: A multi-task recurrent convolutional neural network for forest classification and structural estimation," *Remote Sens.*, vol. 11, no. 7, p. 768, Mar. 2019.

[14] W. Bouachir, K. E. Ihou, H.-E. Gueziri, N. Bouguila, and N. Belanger, "Computer vision system for automatic counting of planting microsites using UAV imagery," *IEEE Access*, vol. 7, pp. 82491–82500, 2019.

[15] H. Yalcin, "Plant phenology recognition using deep learning: Deep-Pheno," in *Proc. 6th Int. Conf. Agro-Geoinform.*, Fairfax, VA, USA, 2017, pp. 1–5.

[16] M. Kosmala, K. Hufkens, and A. D. Richardson, "Integrating camera imagery, crowdsourcing, and deep learning to improve high-frequency automated monitoring of snow at continental-to-global scales," *PLoS ONE*, vol. 13, no. 12, Dec. 2018, Art. no. e0209649.

[17] D. S. Pureswaran, L. De Grandpré, D. Paré, A. Taylor, M. Barrette, H. Morin, J. Régnière, D. D. Kneeshaw, "Climate-induced changes in host tree–insect phenology May drive ecological state-shift in boreal forests," *Ecology*, vol. 96, no. 6, pp. 1480–1491, Jun. 2015.

[18] A. Dutta and A. Zisserman, "The VGG image annotator (VIA)," Apr. 2019, *arXiv:1904.10699*. [Online]. Available: https://arxiv.org/abs/1904.10699

[19] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[20] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É. Duchesnay, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, Oct. 2011.

[21] F. Chollet. *Keras: Deep Learning Library for Theanoand Tensorflow*. Accessed: 2015. [Online]. Available: https://github.com/fchollet/keras

[22] M. Abadi *et al.*, "TensorFlow: A system for large-scale machine learning," in *Proc. 12th Symp. Operating Syst. Design Implement.*, 2016, pp. 265–283.

[23] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 2117–2125.

[24] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 770–778.

[25] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," Dec. 2014, *arXiv:1412.6980*. [Online]. Available: https://arxiv.org/abs/1412.6980

[26] M. Ester, H.-P. Kriegel, S. Jorg, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proc. KDD*, vol. 96, 1996, pp. 226–231.

[27] W. McKinney, "Data structures for statistical computing in python," in *Proc. 9th Python Sci. Conf.*, vol. 445, 2010, pp. 51–56.

[28] T. E. Oliphant, *Guide to Numpy*, vol. 1. New York, NY, USA: Trelgol Publishing, 2006.

[29] G. Bradski, "The opencv library," *Dr Dobb's J. Softw. Tools*, vol. 25, pp. 120–125, 2000.

[30] A. Mayer, "Phenology and citizen science," *BioScience*, vol. 60, no. 3, pp. 172–175, Mar. 2010.

[31] M. D. Schwartz, J. L. Betancourt, and J. F. Weltzin, "From Caprio's lilacs to the USA national phenology network," *Frontiers Ecol. Environ.*, vol. 10, no. 6, pp. 324–327, Aug. 2012.

[32] K. K. Fuccillo, T. M. Crimmins, C. E. De Rivera, and T. S. Elder, "Assessing accuracy in citizen science-based plant phenology monitoring," *Int. J. Biometeorol.*, vol. 59, no. 7, pp. 917–926, Jul. 2015.

[33] C. G. Willis, E. Law, A. C. Williams, B. F. Franzone, R. Bernardos, L. Bruno, C. Hopkins, C. Schorn, E. Weber, D. S. Park, and C. C. Davis, "CrowdCurio: An online crowdsourcing platform to facilitate climate change studies using herbarium specimens," *New Phytol.*, vol. 215, no. 1, pp. 479–488, Jul. 2017.

[34] T. B. Brown, K. R. Hultine, H. Steltzer, E. G. Denny, M. W. Denslow, J. Granados, S. Henderson, D. Moore, S. Nagai, M. Sanclements, A. Sánchez-Azofeifa, O. Sonnentag, D. Tazik, and A. D. Richardson, "Using phenocams to monitor our changing Earth: Toward a global phenocam network," *Frontiers Ecol. Environ.*, vol. 14, no. 2, pp. 84–93, Mar. 2016.

[35] T. F. Keenan, B. Darby, E. Felts, O. Sonnentag, M. A. Friedl, K. Hufkens, J. O'Keefe, S. Klosterman, J. W. Munger, M. Toomey, and A. D. Richardson, "Tracking forest phenology and seasonal physiology using digital repeat photography: A critical assessment," *Ecological Appl.*, vol. 24, no. 6, pp. 1478–1489, Sep. 2014.

[36] O. Sonnentag, K. Hufkens, C. Teshera-Sterne, A. M. Young, M. Friedl, B. H. Braswell, T. Milliman, J. O'Keefe, and A. D. Richardson, "Digital repeat photography for phenological research in forest ecosystems," *Agricult. Forest Meteorol.*, vol. 152, pp. 159–177, Jan. 2012.

[37] A. D. Richardson, B. H. Braswell, D. Y. Hollinger, J. P. Jenkins, and S. V. Ollinger, "Near-surface remote sensing of spatial and temporal variation in canopy phenology," *Ecological Appl.*, vol. 19, no. 6, pp. 1417–1428, Sep. 2009.

[38] T. Mizunuma, M. Wilkinson, E. L. Eaton, M. Mencuccini, J. I. L. Morison, and J. Grace, "The relationship between carbon dioxide uptake and canopy colour from two camera systems in a deciduous forest in southern E ngland," *Funct. Ecol.*, vol. 27, no. 1, pp. 196–207, Feb. 2013.

[39] A. J. Miller-Rushing, T. T. Høye, D. W. Inouye, and E. Post, "The effects of phenological mismatches on demography," *Philos. Trans. Roy. Soc. B, Biol. Sci.*, vol. 365, no. 1555, pp. 3177–3186, Oct. 2010.

[40] A. S. Razavian, H. Azizpour, J. Sullivan, and S. Carlsson, "CNN features off-the-shelf: An astounding baseline for recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, Columbus, OH, USA, Jun. 2014, pp. 512–519.

[41] S. Xie, R. Girshick, P. Dollar, Z. Tu, and K. He, "Aggregated residual transformations for deep neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, Jul. 2017, pp. 5987–5995.

[42] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v4, inception-ResNet and the impact of residual connections on learning," in *Proc. 21st AAAI Conf. Artif. Intell.*, 2017.

[43] Z. Li, C. Peng, G. Yu, X. Zhang, Y. Deng, and J. Sun, "Light-Head R-CNN: In defense of two-stage object detector," Nov. 2017, *arXiv:1711.07264*. [Online]. Available: https://arxiv.org/abs/1711.07264

[44] E. Raczko and B. Zagajewski, "Comparison of support vector machine, random forest and neural network classifiers for tree species classification on airborne hyperspectral APEX images," *Eur. J. Remote Sens.*, vol. 50, no. 1, pp. 144–154, Jan. 2017.

[45] P. Helber, B. Bischke, A. Dengel, and D. Borth, "EuroSAT: A novel dataset and deep learning benchmark for land use and land cover classification," Aug. 2017, *arXiv:1709.00029*. [Online]. Available: https://arxiv.org/abs/1709.00029

DAVID L. P. CORREIA received the Ph.D. degree in forest science from Laval University, Québec, Canada, in 2018. He is currently a Research Scientist in forest ecology and climate change with the Canadian Forest Service. His work focuses mainly on the application of artificial intelligence and computer vision techniques to the study of forest ecology. He is also studying the impacts of climate change on land cover change in the Canadian northern protected areas.

WASSIM BOUACHIR received the M.Sc. degree in computer science from the Université de Moncton, New-Brunswick, Canada, and the Ph.D. degree in computer engineering from Polytechnique Montréal, Québec, Canada, in 2015. Following his doctoral studies in computer vision, he joined the École de Technologie Supérieure, Québec, Canada, as a Postdoctoral Researcher. He is currently a Professor of computer science with TÉLUQ University. His research interests include image analysis, video processing, and machine learning. His research works aim to develop intelligent computer vision systems for several application areas, such as security, environment sciences, and health-care applications.

DAVID GERVAIS received the M.Sc. degree in biology (forest entomology) from Concordia University, Montreal, Canada. He currently works with several research teams (climate change and post-disturbance forest dynamics) with the Canadian Forest Service specializing in data management, dendrochronology, GIS, and the coordination of field work logistics.

DEEPA PURESWARAN received the Ph.D. degree in forest entomology from Simon Fraser University, in 2004. She did her postdoctoral work at the Dartmouth College, New Hampshire, and at Michigan State University. She has been with the Canadian Forest Service, since 2008. She was involved in the study of outbreak ecology of native and non-native forest insects and ecosystem response following disturbances under climate change.

DANIEL D. KNEESHAW has been a Professor at the Department of Biological Sciences, Université du Québec à Montréal, since 2001. His work focuses on the development of sustainable forest management indicators based on a better understanding of forest dynamics. He is also working on the development of stand-level and landscape-level forest dynamics modeling tools.

LOUIS DE GRANDPRÉ is currently a Research Scientist in forest ecology with Natural Resources Canada. His work focuses on the impact of disturbances on boreal forest dynamics. He is also involved in study of effect of climate warming on plant phenology and potential phenological mismatch between plant host and pests.